Lec 11 MATH 343

$X \in \mathbb{R}^{n \times (p+1)}$ design matrix of covariates

$\beta \in \mathbb{R}^{p+1}$ linear slope parameters ~~defd~~ defined by $h^*(\vec{x}) = \vec{x}\vec{\beta}$

best linear approx to $f(\vec{x})$

$\vec{\varepsilon} \in \mathbb{R}^n$ errors

$\vec{y} \in \mathbb{R}^n$ responses $\Rightarrow \vec{y} = X\vec{\beta} + \vec{\varepsilon}$

$\vec{b} = (X^\top X)^{-1} X^\top \vec{y}$ OLS estimate of $\vec{\beta}$, $H = X(X^\top X)^{-1}X^\top$ projection matrix onto colsp$[X]$.

$\hat{\vec{y}} = X\vec{b} = H\vec{y}$ OLS in-sample predictions

$\vec{e} = \vec{y} - \hat{\vec{y}} = (I - H)\vec{y}$ OLS in-sample residuals

342 ↑

$\vec{y} = X\vec{b} + \vec{e}$

343 ↓

Assume $\vec{\varepsilon}$ is a draw from $\vec{\mathcal{E}} \sim N_n(\vec{0}, \sigma^2 I_n)$

$\vec{Y} = X\vec{\beta} + \vec{\mathcal{E}} \Rightarrow \vec{Y}$ is random

$\vec{B} \sim N_{p+1}(\vec{\beta}, \sigma^2 (X^\top X)^{-1})$

$\hat{\vec{Y}} = X\vec{B} = H\vec{Y} \sim N_n(X\beta, \sigma^2 H)$

$E = \vec{Y} - \hat{\vec{Y}} = (I-H)\vec{Y} \sim N_n(\vec{0}_n, \sigma^2 (I - H))$

Under overfitting $\Rightarrow p = n+1 \Rightarrow H = I \underset{\Rightarrow I - H = 0_{n \times n}}{\Rightarrow} \vec{E} \sim N_n(\vec{0}_n, 0_{n\times n}) = \text{Deg}(\vec{0}_n) \Rightarrow \vec{Y} = \hat{\vec{Y}} \sim N(X\beta, \sigma^2 I_n)$

By Cochran's Thm, $\frac{1}{\sigma^2}\|\vec{E}\|^2 \sim \chi^2_{n-(p+1)} \overset{340}{=} \text{Gamma}\left(\frac{n-(p+1)}{2}, \frac{1}{2}\right) \overset{340}{\Rightarrow} \|\vec{E}\|^2 \sim \text{Gamma}\left(\frac{n-(p+1)}{2}, \frac{1}{2\sigma^2}\right)$

$\Rightarrow E\left[\|\vec{E}\|^2\right] = \frac{\frac{n-(p+1)}{2}}{\frac{1}{2\sigma^2}} = \sigma^2(n-(p+1)) \Rightarrow \frac{E\left[\|\vec{E}\|^2\right]}{n-(p+1)} = \sigma^2 \Rightarrow \frac{SSE}{n-(p+1)}$ is an unbiased estimate of σ^2

Define $R^2_{adj} := 1 - \dfrac{\frac{SSE}{n-(p+1)}}{\frac{SST}{n-1}}$ (unbiased estimate of $Var[\mathcal{E}]$; unbiased estimate of $Var[Y]$) $= 1 - \underbrace{\frac{n-1}{n-(p+1)}}_{c \,\leftarrow\, \text{correction factor}} \frac{SSE}{SST}$

$MSE := \frac{SSE}{n-(p+1)}$

if $p \uparrow \Rightarrow c \uparrow \Rightarrow R^2_{adj} \downarrow$ more honest

$\Rightarrow SSE \downarrow$

Previously $R^2 = 1 - \frac{SSE}{SST} \leftarrow$ if $p\uparrow \Rightarrow SSE \downarrow \Rightarrow R^2 \uparrow$

$= 1 - \dfrac{\frac{SSE}{n-1}}{\frac{SST}{n-1}}$ **biased** est. of $Var[\mathcal{E}]$; unbiased est. of $Var[Y]$ not honest!

By Cochran's Thm, $\frac{1}{\sigma^2}\|X(\vec{B}-\vec{\beta})\|^2 \sim \chi^2_{p+1}$ and indep. of $\vec{E}$

This with the previous results gives us

$$\frac{B_j-\beta_j}{s_e\sqrt{(X^TX)^{-1}_{jj}}} \sim T_{n-(p+1)} \quad \overset{3\&D}{\Longrightarrow} \quad \frac{(B_j-\beta_j)^2}{s_e^2\,(X^TX)^{-1}_{jj}} \sim F_{1,\,n-(p+1)}$$

F-t sq. equivalence

And letting $\mu_\alpha = \vec{x}_\alpha \vec{\beta}$, $\quad \frac{\hat{Y}_\alpha - \mu_\alpha}{s_e\sqrt{\vec{x}_\alpha (X^TX)^{-1}\vec{x}_\alpha^T}} \sim T_{n-(p+1)}$

What if we want inference for the new response itself, $\theta = y_*$

Recall our assumption: $\mathcal{E}_1, \ldots, \mathcal{E}_n \overset{iid}{\sim} N(0,1)$

Also assume the new noises are $\mathcal{E}_* \sim N(0,1)$ and indep of $\mathcal{E}_1, \ldots, \mathcal{E}_n$

Consider: $$Y_* - \hat{Y}_* = Y_* - \vec{x}_* \vec{B} = \underbrace{\vec{x}_* \vec{\beta}}_{\text{const}} + \underbrace{\mathcal{E}_*}_{N(0,\sigma^2)} - \underbrace{x_* \vec{B}}_{N(\vec{x}_*\vec{\beta},\, \sigma^2 \vec{x}_* (X^\top X)^{-1} \vec{x}_*^\top)}$$

Due to assumption $\mathcal{E}_*$ and $\vec{B}$ are independent since $\vec{B} = (X^\top X)^{-1} X^\top \vec{Y} = (X^\top X)^{-1} X^\top (X\vec{\beta} + \vec{\mathcal{E}})$

and $\vec{\mathcal{E}}$ is independent of $\mathcal{E}_*$.

$$\Rightarrow Y_* - \hat{Y}_* = N\left(0, \sigma^2 + \sigma^2 \vec{x}_* (X^\top X)^{-1} \vec{x}_*^\top\right) = N\left(0, \sigma^2 (1 + \vec{x}_* (X^\top X)^{-1} \vec{x}_*^\top)\right)$$

$$\Rightarrow \frac{Y_* - \hat{Y}_*}{\sqrt{\sigma^2 (1 + \vec{x}_* (X^\top X)^{-1} \vec{x}_*^\top)}} \sim N(0,1)$$

But σ^2 is unknown! So consider

$$\frac{\left.\dfrac{Y_* - \hat{Y}_*}{\sqrt{\sigma^2 (1 + \vec{x}_* (X^\top X)^{-1} \vec{x}_*^\top)}}\right\} N(0,1)}{\left.\sqrt{\dfrac{\frac{\|\vec{E}\|^2}{\sigma^2}}{n-(p+1)}}\right\} \chi^2_{n-p}} \sim T_{n-(p+1)}$$

indep since $\vec{B}, \vec{E}$ indep.

$$\Rightarrow \frac{\dfrac{Y_* - \hat{Y}_*}{\sqrt{(1 + \vec{x}_* (X^\top X)^{-1} \vec{x}_*^\top)}}}{\sqrt{\dfrac{\|\vec{E}\|^2}{n-(p+1)}}} \sim T_{n-(p+1)}$$

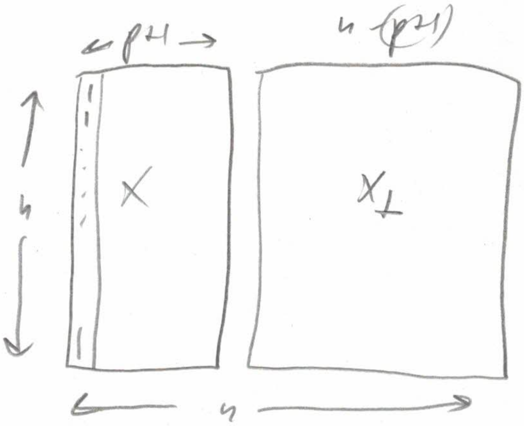
the test statistic for $H_a: Y_* \neq \theta$ is:

$$\frac{\theta - \hat{y}_*}{s_e \sqrt{1 + \vec{x}_* (X^\top X)^{-1} \vec{x}_*^\top}} \in \left[\pm t_{1-\frac{\alpha}{2}, n-(p+1)}\right]$$

$\Rightarrow$ Retain H_0

Which can be inverted for a CI:

$$CI_{y_*, 1-\alpha} = \left[\hat{y}_* \pm t_{1-\frac{\alpha}{2}, n-(p+1)}\, s_e \sqrt{1 + \vec{x}_* (X^\top X)^{-1} \vec{x}_*^\top}\right]$$



Both $X, X_\perp$ are full rank

$\Rightarrow \text{rank}\left[[X \,\vdots\, X_\perp]\right] = n$

$\text{colsp}\left[[X \,\vdots\, X_\perp]\right] = \mathbb{R}^n$

$\text{proj}_X(\vec{y}) = H\vec{y}, \ \text{proj}_{X_\perp}(\vec{y}) = (I - H)\vec{y}$

Just like we can partition $\mathbb{R}^n$ into $\text{colsp}(X)$, $\text{colsp}[X_\perp]$,

we can partition $\text{colsp}[X]$ into $\text{colsp}[\vec{1}_n]$, $\text{colsp}[X_{-1}]$

let $H_1 := \vec{1}_n(\vec{1}_n^\top \vec{1}_n)^{-1}\vec{1}_n = \frac{1}{n}J_n$ ← matrix of all 1's $\qquad H_1\vec{y} = \bar{y}\vec{1}_n$

let $H_{-1} := H - H_1 = X(X^\top X)^{-1}X^\top - \frac{1}{n}J_n$, the projection onto $\text{colsp}[X_{-1}]$ which is orthogonal to $\vec{1}_n$.

$$\hat{\vec{y}} = H\vec{y} = \left(H_1 + (H - H_1)\right)\vec{y} = H_1\vec{y} + (H - H_1)\vec{y}$$

$$(H_1\vec{y})^\top((H - H_1)\vec{y}) = (\bar{y}\vec{1}_n)^\top(H\vec{y} - H_1\vec{y}) = \bar{y}\,\vec{1}_n^\top(H\vec{y} - \bar{y}\vec{1}_n) = \bar{y}\,\vec{1}_n^\top H^\top\vec{y} - \bar{y}^2\,\vec{1}_n^\top\vec{1}_n$$
$$= \bar{y}\,(H\vec{1})^\top\vec{y} - n\bar{y}^2 = \bar{y}\,\vec{1}^\top\vec{y} - n\bar{y}^2 = \bar{y}\sum y_i - n\bar{y}^2 = n\bar{y}^2 - n\bar{y}^2 = 0$$

$$\text{rank}[H_1] = 1, \quad \text{rank}[H_{-1}] = tr[H - H_1] = (p+1) - (1) = p$$

Let's use Cochran's Thm again:

$$\frac{1}{\sigma^2}\vec{\varepsilon}^\top\vec{\varepsilon} = \frac{1}{\sigma^2}\vec{\varepsilon}^\top H_1\vec{\varepsilon} + \frac{1}{\sigma^2}\vec{\varepsilon}^\top(H - H_1)\vec{\varepsilon} + \frac{1}{\sigma^2}\vec{\varepsilon}^\top(I - H)\vec{\varepsilon} \sim \chi^2_n$$

$$\Rightarrow \frac{1}{\sigma^2}\vec{\varepsilon}^\top H_1\vec{\varepsilon} \sim \chi^2_1 \text{ indep of } \frac{1}{\sigma^2}\vec{\varepsilon}^\top(H - H_1)\vec{\varepsilon} \sim \chi^2_p \text{ indep of } \frac{1}{\sigma^2}\vec{\varepsilon}^\top(I - H)\vec{\varepsilon} \sim \chi^2_{n-(p+1)}$$

$$\frac{1}{\sigma^2}\vec{\varepsilon}^\top(H - H_1)\vec{\varepsilon} = \frac{1}{\sigma^2}\vec{\varepsilon}^\top(H - H_1)^\top(H - H_1)\vec{\varepsilon} = \frac{1}{\sigma^2}\|(H - H_1)\vec{\varepsilon}\|^2 = \frac{1}{\sigma^2}\|(H - H_1)(\vec{Y} - X\vec{\beta})\|^2$$
$$= \frac{1}{\sigma^2}\|HY - HX\beta - H_1\vec{Y} + H_1X\vec{\beta}\|^2 = \frac{1}{\sigma^2}\|\hat{\vec{Y}} - X\beta - \bar{y}\vec{1}_n + H_1X\vec{\beta}\|^2$$

From Cochran's Thm:

$$\frac{\frac{\frac{1}{\sigma^2}\vec{\varepsilon}^\top(H - H_1)\vec{\varepsilon}}{p}}{\frac{\frac{1}{\sigma^2}\vec{\varepsilon}^\top(I - H)\vec{\varepsilon}}{n - p + 1}} \sim F_{p, n-(p+1)} \quad \Rightarrow \quad \frac{\frac{\|\hat{\vec{Y}} - X\beta - \bar{y}\vec{1}_n + H_1X\vec{\beta}\|^2}{p}}{\frac{\|\vec{E}\|^2}{n - (p+1)}} \sim F_{p, n-(p+1)}$$

supports $(0, \infty)$